УДК 004.855

РЕАЛИЗАЦИЯ РАЗНОВИДНОСТЕЙ ОТКРЫТОЙ ФОРМЫ КОМПЬЮТЕРНОГО ТЕСТИРОВАНИЯ ЗНАНИЙ

В.Б. Немировский, А.К. Стоянов

Томский политехнический университет
E-mail: stoj-ak@ad.cctpu.tpu.edu.ru, nemir-vb@ad.cctpu.tpu.edu.ru

Описывается реализация систем открытой формы компьютерного тестирования знаний для проверки знаний по основам программирования и тем учебным дисциплинам, в которых вопросы требуют формулировки пространного ответа. Рассматриваются особенности реализации систем тестирования, их состав, этапы работы с ними.

В настоящее время существует большое количество компьютерных систем тестирования знаний. Они охватывают разные области знания, предоставляют тем, кто этими системами пользуется, разнообразные средства достижения поставленных целей. Составителям тестов предоставляется возможность контролировать знания, а пользователям тестов – проверить знания.

Согласно одной из разновидностей классификации тестов, основанной на характере ответов на вопросы, существует два вида тестов – тесты закрытого типа и открытого типа. В тестах закрытого типа ответы предписаны, и испытуемому необходимо лишь выбрать из предложенных вариантов ответов тот или иной вариант. В тестах открытого типа испытуемому необходимо самостоятельно сформулировать ответ, дописав слово, словосочетание, предложение, знак, формулу и т. д. [1].

Все разнообразие компьютерных систем тестирования объединяет то, что они проверяют, в основном, лишь начальный уровень усвоения знаний – репродуктивное узнавание. Это связано с тем, что большая часть компьютерных систем тестирования знаний основана на разновидностях одной формы тестирования – закрытой. Эта форма тестирования легко реализуема, чем и обусловлено ее широкое распространение.

Существует много исследований, посвященных эффективности закрытой формы тестирования с выбором ответов. В [2], например, отмечено, что контроль знаний, построенный на основе выборочных ответов, принципиально не может выявить более высокого уровня обучения, чем уровень «знакомства». Другие виды закрытой формы тестирования (альтернативные ответы, соответствие, последовательность, множественный выбор, исключение лишнего, аналогии) в том или ином виде также реализуют выбор из предложенных вариантов, и, следовательно, также позволяют проверять лишь начальный уровень приобретения знаний.

Следует отметить, что даже в тех случаях, когда характер поставленного вопроса позволяет проверять более высокую степень приобретения знаний, например, понимание или применение, использование закрытой формы тестирования приводит к тому, что снижается его диагностическая эффективность. В качестве иллюстрации сказанному можно привести вопрос, проверяющий умение применять полученные знания, который взят из документации к системе тестирования OpenTest [3].

Даны три цифры (например, 1, 2, 3). Сколько трехзначных чисел можно составить из них, если цифры не повторяются?

- 3;
- 6;
- 9;
- 12.

Очевидно, что приведение нескольких заданных ответов снижает эффективность тестирования. В такой форме резко повышается вероятность случайного выбора правильного ответа при полном неумении решить задачу. В то же время, изменив форму тестирования с закрытой на открытую форму, т. е. убрав варианты ответов, действительно можно проверить умение отвечающего решать подобные задачи.

Из всего сказанного можно сделать вывод, что, не отказываясь полностью от использования заданий закрытого типа, разработчики автоматизированных систем контроля знаний должны думать о расширении разнообразия видов тестирования. Иными словами, если требуется проверять умение понимать, а не узнавать предъявленный материал, умение применять его для получения ответа на вопрос, потребуется наличие систем, реализующих открытую форму тестирования.

В тестах открытого типа выделяют следующие возможные варианты ответов:

- задания дополнения – испытуемый должен сам сформулировать ответы с учетом предусмотренных в задании ограничений;
- свободного изложения – испытуемый должен самостоятельно сформулировать ответы, ибо никакие ограничения на них в задании не накладываются.

Известно, что в настоящее время задача компьютерной реализации открытой формы тестирования в общем случае (свободно-конструируемый развернутый ответ на естественном языке) не решена. Причина заключается в сложности извлечения смысла из фраз слабо формализованного естественного языка, на котором формулируется ответ. При этом нужно отметить, что возможна простейшая реализация открытой формы тестирования в случае, когда результатом ответа на вопрос является число или фиксированная строка. В этом случае проверка правильности введенного ответа осуществляется сравнением с эталоном. Однако для ситуации, в которой требуется развернутый ответ, такой подход не эффективен. Например, для тестирования знаний основ программирования ответом на некоторые вопросы тестов может быть исходный код программы и результат ее выполнения. Другая ситуация возникает при необходимости дать развернутый ответ, описывающий некоторый алгоритм, либо последовательность действий, либо состав некоторой системы и т. п.

Мы рассмотрим вопросы реализации двух систем тестирования, соответствующих описанным выше ситуациям.

Система тестирования знаний основ программирования

При обучении основам программирования контроль приобретенных практических умений и способности создавать правильно работающую программу намного полезнее для обучения, чем ответы на общие вопросы в тестирующем диалоге. Именно в этой ситуации требуется давать развернутые ответы в виде текста программ. Проверка правильности ответа должна включать стадии синтаксической и семантической правильности написанной программы. Основная идея реализации системы связана с тем, что предметная область включает в себя проблемно-ориентированный язык – язык программирования. Синтаксическую проверку текста на этом языке можно передать компилятору выбранного языка программирования. А семантическая проверка осуществляется тестированием синтаксически правильной программы при заданном наборе исходных данных.

Разработанная система предназначена не только для тестирования знаний основ программирования, но и для итогового контроля этих знаний в форме зачета или экзамена. Важно отметить, что система предназначена для студентов, изучающих программирование как общеобразовательную дисциплину. Поэтому система должна избавлять студента от необходимости знать тонкости и детали реализации языка программирования, используемые библиотеки и особенности конкретной среды разработки, конкретную реализацию ввода и вывода. Кроме того, система предоставляет средства подготовки материалов для тестирования (инструмент преподавателя). Одним из главных принципов организации системы автоматизированного тестирования является возможность преподавателя самостоятельно или через оператора вносить задания в систему, не прибегая к услугам программиста.

При разработке первого компонента системы, осуществляющего процесс тестирования студентов, был проведен предварительный анализ вопросов, встречающихся в зачетных и экзаменационных билетах по программированию. В результате выделено три основные категории, требующие ответа в виде:

- значения (числа или строки);
- исходного кода программы;
- исходного кода процедуры или функции.

Разработанная система тестирования выполняет следующие основные задачи:

- предоставляет возможность случайного выбора вопросов (билетов) для тестирования;
- предоставляет средства для ввода значений ответов на вопросы первой категории и текста программы для остальных категорий вопросов;
- на основе введенного текста программы генерирует максимально близкий к нему исходный код, дополненный формализованным вводом и выводом и средствами организации контроля (например, средствами записи полученного результата в файл определенной структуры для последующего сравнения с эталоном ответа);

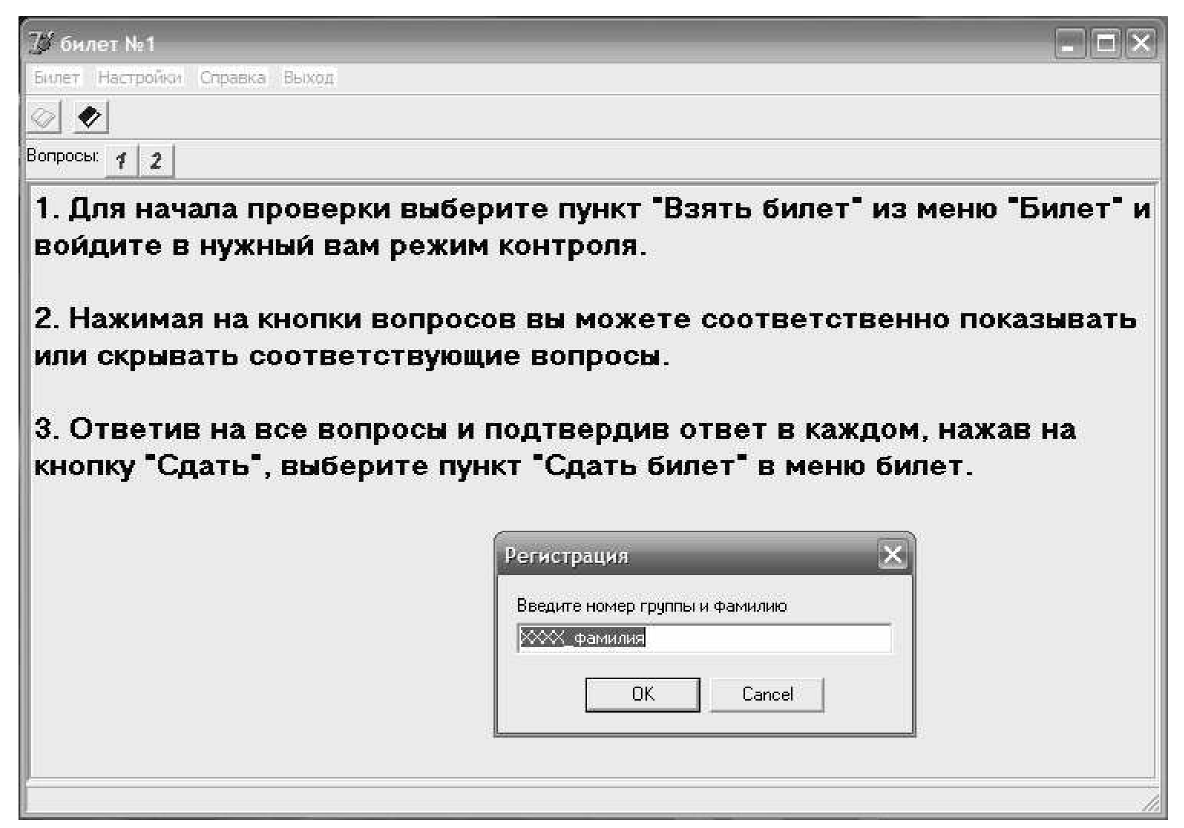


Рис. 1. Начальное окно и окно регистрации системы тестирования знаний по основам программирования

- вызывает компилятор для проверки синтаксической правильности программы;
- организует вывод сообщений компилятора в окно с текстом программы студента;
- учитывает количество неуспешных компиляций;
- организует запуск успешно откомпилированной программы в консольном режиме для ее отладки;
- предоставляет тестовые наборы данных для семантического тестирования созданной программы;
- производит сравнение эталонных и полученных результатов и оценивает результат тестирования.

В качестве языка программирования для студентов выбран язык Паскаль, который используется при изучении основ программирования для непрофильных специальностей. Система генерирует код в виде проекта Delphi для компилятора фирмы Borland, включенного в поставку Delphi (выбор языка и компилятора принципиального значения не имеет).

Несколько слов о втором компоненте системы – инструменте преподавателя. Он реализован в виде текстового редактора, сохраняющего подготовленный преподавателем файл наборов вопросов (билетов) и эталоны ответов к ним в зашифрованном виде, чтобы студенту было невозможно узнать его содержимое.

Здесь представлены несколько скриншотов, иллюстрирующих основные этапы работы системы тестирования. После запуска программы студент должен выбрать форму тестирования (зачет или экзамен), пройти регистрацию и получить доступ к вопросам билета, выбранного системой случайным образом. На рис. 1 представлено начальное окно системы, а на рис. 2 и 3 – окна диалогов для ответов на вопросы билета. При ответе на вопрос, требующий ввода текста программы щелчок по кнопке «Сдать» вызывает появление окна (рис. 4), в котором приведены исходные данные, необходимые для тестирования семантической правильности программы.

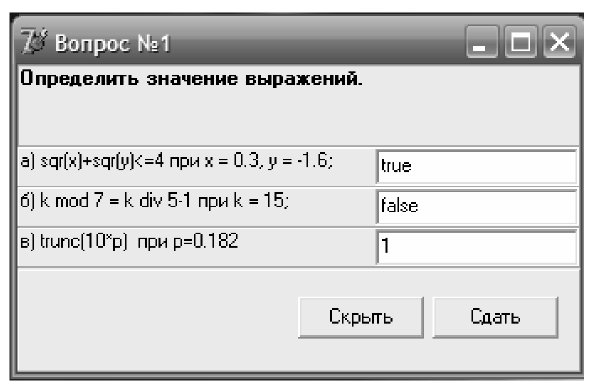


Рис. 2. Диалог с полями для ввода ответов на вопросы первой категории

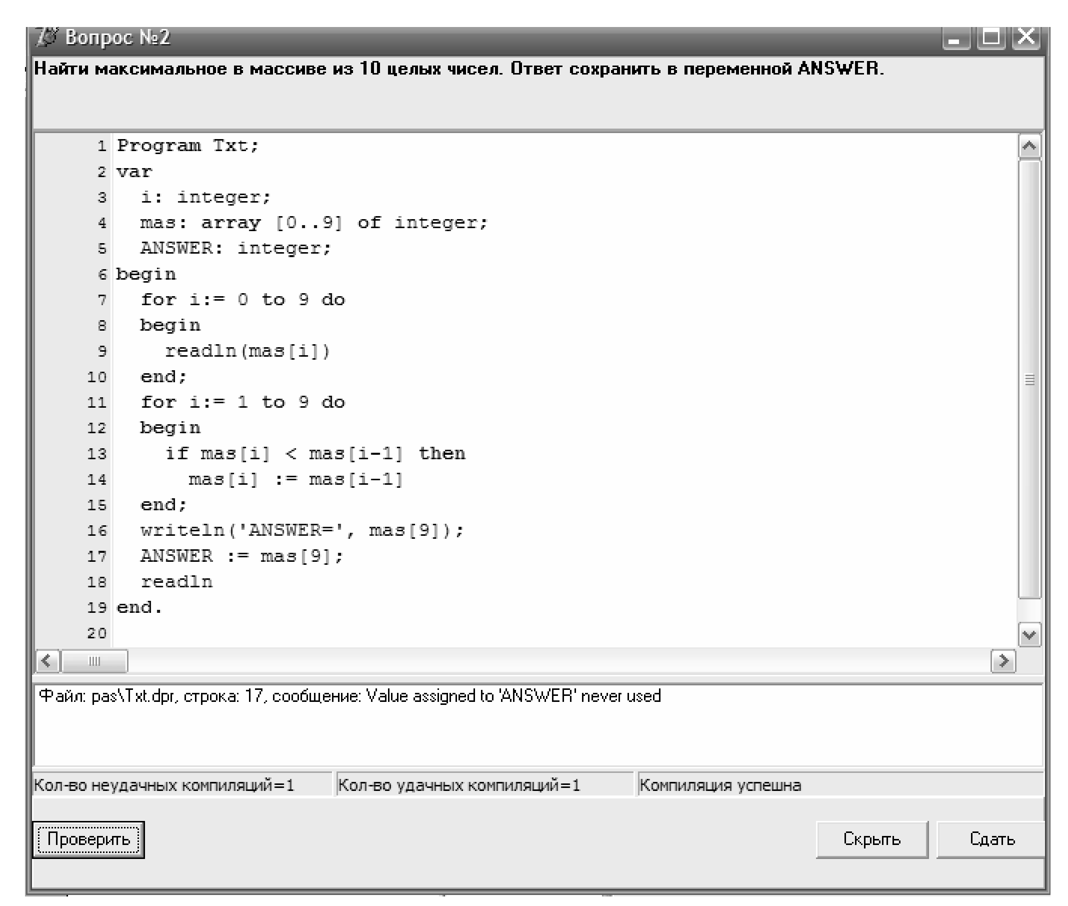


Рис. 3. *Окно для ответа на вопросы второй категории с текстом программы и сообщениями компилятора*

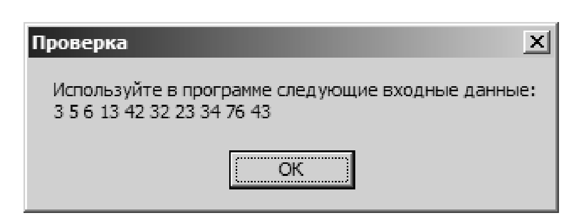


Рис. 4. *Окно сообщения с исходными данными для тестирования семантической правильности программы*

Система контроля, основанная на открытой форме тестирования с использованием конструируемого ответа

Разработка системы была обусловлена необходимостью автоматизировать проверку знаний начального уровня освоения компьютера, а также программных средств информационных технологий. Особенность ответов при проверке этой тематики в том, что: 1) они редко укладываются в одно слово или число; 2) часто включают перечень последовательности действий, необходимых для достижения некоторой цели, или же перечень элементов изучаемой системы. Применение закрытой формы тестирования, как указывалось выше, не позволяет должным образом проверить полученные знания. Использование открытой формы тестирования в простейшем виде также малоэффективно из-за отмеченной пространности ответов. Для повышения эффективности открытой формы тестирования в данном случае было решено использовать её разновидность, основанную на выборочно-конструируемых ответах.

В ответах этого типа отвечающий должен собирать правильный ответ из предложенных элементов ответа или соотносить их между собой. Элементами ответа могут выступать предложение, одна или несколько фраз, числа, совокупности формул [5, 6] или слов, или графиков и рисунков [7]. Известные нам реализации этой формы тестирования обладают некоторыми общими признаками. Во-первых, элементы ответа (и сам ответ) не являются фразами естественного языка. Они представляются в виде фрагментов химических формул или рисунков, отображающих некоторое понятие. Во-вторых, элементы ответа нумеруются, а сам ответ представляется числовой последовательностью. Для учебных дисциплин, в которых ответ на вопрос формулируется на естественном языке, такое числовое его представление снижает обучающую ценность вопроса.

Чтобы использовать выборочно-конструируемые ответы в нашем случае, мы остановились на варианте формулировки ответа на естественном языке. Сформулированный правильный ответ при подготовке теста разбивается на набор отдельных элементов – фрагментов, каждый из которых мо-

жет состоять из нескольких слов. Фрагмент может быть логически завершенным, но это не обязательно. Для уменьшения вероятности угадывания в набор включаются также фрагменты других ответов-дистракторов. Фрагменты в наборе располагаются случайным образом и предъявляются для ответа. Сконструированный из фрагментов ответ сравнивается с правильным ответом.

При таком подходе появляется вероятность угадывания ответа за счет смысловой несогласованности фрагментов правильного ответа и ответов-дистракторов. Для уменьшения такой вероятности все вопросы и фрагменты ответов к ним в процессе создания теста классифицируются по отдельным видам знаний предметной области тестирования. Для классификации используются категории объектно-структурного анализа, предназначенного для практического структурирования знаний [4]. Согласно этому анализу, знания любой предметной области могут быть разделены на 8 категорий:

- ЗАЧЕМ-знания;
- КТО-знания;
- ЧТО-знания;
- КАК-знания;
- ГДЕ-знания;
- КОГДА-знания;
- ПОЧЕМУ-знания;
- СКОЛЬКО-знания.

При предъявлении списка фрагментов ответов испытуемому все фрагменты берутся из ответов для вопросов одной категории. Это позволяет обеспечить смысловую согласованность фрагментов ответов на эти вопросы.

Система состоит из двух подсистем: подсистемы разработчика теста – преподавателя и подсистемы пользователя теста – студента. Подсистема преподавателя содержит программные средства, с помощью которых создаются и редактируются вопросы и ответы теста, выполняется их классификация. Подсистема студента включает в себя средства выбора и отображения вопросов и фрагментов ответов на них, средства оценки выполненного теста.

На рис. 5 приведен скриншот окна программной части системы КИТ (Компьютерный Инструментарий для Тестирования), предназначенной для создания вопросов и элементов ответов на них на стадии классификации (подсистема преподавателя). Из полного списка вопросов выбираются вопросы и фрагменты ответов, соответствующие категории ЧТО-знания.

На рис. 6 приведен скриншот окна программной части системы КИТ – подсистема студента. Предлагаемые для ответа фрагменты отображаются в левой части окна в виде списка. Из списка требуется выбрать правильные фрагменты и расположить в нужной последовательности в предназначенной для ответа области.

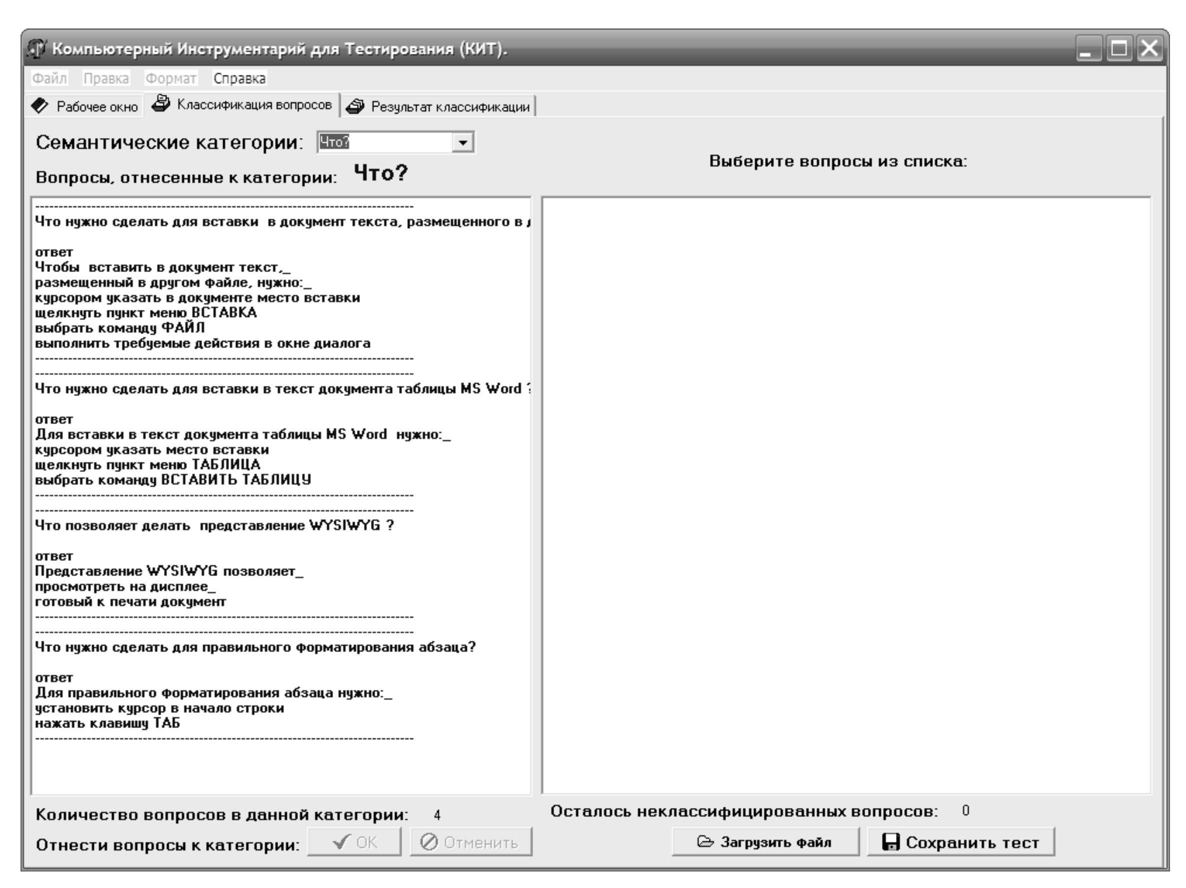


Рис. 5. *Окно системы КИТ на этапе классификации вопросов*

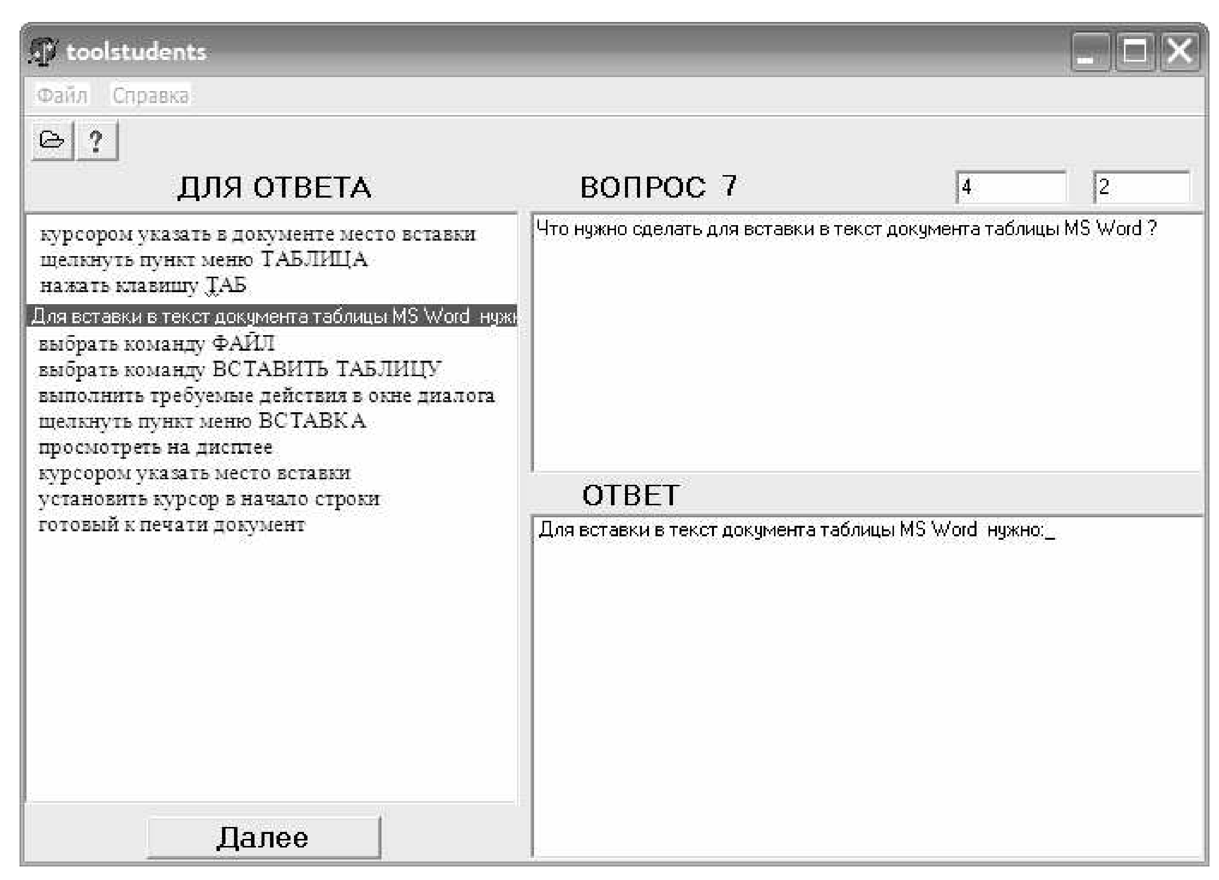


Рис. 6. *Окно системы КИТ на этапе тестирования*

Рассматриваемая реализация дает возможность проверять наличие знаний, необходимых для понимания сути поставленного вопроса, а также ответа на него. Значительно уменьшается и вероятность случайного угадывания ответа. Действительно, пусть правильный ответ состоит из k фрагментов, которые выбираются из списка, содержащего n фрагментов. Тогда число всевозможных комбинаций из k фрагментов с учетом их следования равно числу размещений из n по k, т. е. $A_n^k=n!/(n-k)!$, и, следовательно, вероятность случайного угадывания ответа пропорциональна величине $p_1=1/A_n^k$. Для сравнения, вероятность угадывания при выборе правильного ответа из нескольких предложенных составляет $p_2=1/m$, где m – количество предлагаемых вариантов ответа [3]. Нетрудно оценить выигрыш, получаемый в нашем случае. Пусть число предлагаемых вариантов ответа для закрытого тестирования равно полному числу фрагментов (эквивалентная ситуация выбора), т. е. $m=n$. Тогда отношение $p_2/p_1=(1/m)A_n^k$. Так, для $m=n=6$ и $k=3$ величина этого отношения равна 20, т. е. вероятность случайного угадывания ответа при открытом тестировании с конструированием ответа в 20 раз меньше, чем при закрытом тестировании с выбором одного правильного ответа из нескольких предложенных.

При создании тестов с помощью систем описанного типа важным моментом является правильная классификация вопросов по категориям знаний. Это предъявляет соответствующие требования и к самим вопросам. Иногда для лучшей смысловой согласованности фрагментов ответов приходится перефразировать вопрос таким образом, чтобы он из одной категории знаний перешел в другую.

Следующее замечание связано делением ответа на отдельные фрагменты. Лучше всего фрагментации поддаются ответы на вопросы, включающие в себя списки. В этом случае фрагментом является отдельный элемент списка. Чаше вопросы такого рода относятся к категориям КАК-знания и ЧТО-знания. Они требуют ответов, содержащих либо перечень действий, либо перечень элементов какой-либо системы. Что касается вопросов, относящихся к другим категориям знаний, то в ответах на них фрагментами лучше выбирать логически завершенные части всей фразы, содержащей ответ.

Описанная система может быть также использована для контроля знаний не только в упомянутом, но и в любых подобных случаях, когда на вопросы требуются ответы в форме перечисления.

Таким образом, описанные системы тестирования показывают возможность полноценной реализации открытой формы тестирования:

- со свободно-конструируемым ответом на основе проблемно-ориентированного языка;
- с выборочно-конструируемым ответом на естественном языке.

Описанные системы позволяют проверять более высокие уровни знания, чем репродуктивное узнавание. Их встраивание в компьютерные обучающие системы, связанные с изучением соответствующих дисциплин, значительно повышает образовательные возможности таких систем.

СПИСОК ЛИТЕРАТУРЫ

1. http://www.testbox.softvea.ru/docs/theory/classification.html
2. Баранов В.Ю. Формы ответов при автоматизированном контроле знаний студентов // Науковий вісник ПДПУ ім. К.Д. Ушинського. – 2004. – № 7. – С. 17–24.
3. http://www.opentest.com.ua/
4. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. – СПб.: Питер, 2000. – 384 с.
5. Щербаков В.В., Капустин Ю.М., Федосеев А.С. Разработка электронных учебных изданий. – М.: РХТУ им. Д.И. Менделеева, 2002. – 84 с.
6. Щербаков В.В., Мамакина Е.А., Барботина Н.Н. Разработка компьютерных аттестационных педагогических измерительных материалов по химии // Современные информационные технологии: Труды Междунар. научно-техн. конф. – Пенза, 2002, – С. 65–66.
7. http://window.edu.ru/window_catalog/files/6773/2000_2_070

УДК 681.3.01